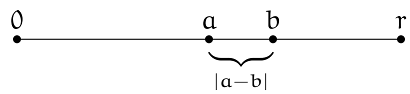
Math Club Notes: 2005 May 16

1 Geometric vs algebraic proofs

Before Eileen and Ray arrived, Mehran and I discussed the relative advantages of geometric and algebraic proofs. For example, consider this small theorem:

> If $0 \le a \le r$ and $0 \le b \le r$, then $0 \le |a - b| \le r$.

Here's a geometric proof:



Obviously, the distance between a and b (which is what $|a - b|$ means) cannot be greater than r.

Here's an algebraic proof:

$$\left.\begin{array}{l}\left.\begin{array}{rcl} a \le r & \implies & a - b \le r - b \\ 0 \le b & \implies & r - b \le r \end{array}\right\} \implies a - b \le r \\ \left.\begin{array}{rcl} b \le r & \implies & b - a \le r - a \\ 0 \le a & \implies & r - a \le r \end{array}\right\} \implies b - a \le r \end{array}\right\} \implies |a - b| \le r$$

The other half, $0 \le |a - b|$, is true by definition.

The geometric proof feels very clear; it makes the theorem "obvious". In this use, "obvious" means "intuitively obvious", that is, our geometric intuition tells us without analysis that the theorem is true.

Unfortunately, intuition (pretty much by definition) does not concern itself with explanations. I find it quite difficult to look at this geometric proof and state exactly what assumptions I'm making when I judge, on the basis of this picture, that the theorem is true.

And anyone who's done a reasonable amount of geometry will have had the experience of realizing that a picture they'd drawn tacitly included a false assumption, or an assumption which was not given in the hypothesis.

In the algebraic proof, on the other hand, the assumptions are all explicit. They're just the axioms of order from the previous note (and a couple small theorems derived therefrom).

This is (one reason) why we bother with abstract algebra. By making our assumptions about numbers explicit as axioms, and refusing to appeal to geometric intuition, we make our proofs more sound. We also discover exactly what the logical relationships actually are among the large body of intuitively obvious facts.

(We also discover that some of those axioms apply to objects which are not numbers, and for which our intuition may not serve us well.)

2 nth roots

We looked at two proofs of the existence and uniqueness of nth roots. To be precise, the theorem in question is this:

> For any positive real number α and positive integer n, there exists exactly one positive real number c such that $c^n = \alpha$.

Uniqueness is the easy part: if $0 < c_1 < c_2$, then (by induction on n) $c_1^n < c_2^n$ for all positive integers n. So a positive number can have at most one nth root. The trick is proving existence.

2.1 Proof by IVT

The first proof we looked at uses a fundamental result of calculus, namely the Intermediate Value Theorem (hereafter "IVT"), which states:

> If f is continuous on $[a, b]$, and $f(a) < 0 < f(b)$, then there exists a real number $c \in (a, b)$ such that $f(c) = 0$.

(Paraphrased: if a continuous function crosses the x-axis, then it crosses the x-axis.) This theorem in hand, we proceed to the proof.

Let α be a positive real number, and let n be a positive integer. We wish to show that there exists a positive number c such that $c^n = \alpha$.

Consider the function $f(x) = x^n - \alpha$. Note that $f(0) = 0^n - \alpha = -\alpha < 0$, since by hypothesis $\alpha > 0$. We also have $0 < f(\alpha + 1)$; for

$$\begin{aligned} f(\alpha+1) &= (\alpha+1)^n - \alpha \\ &= \sum_{k=0}^{n} \binom{n}{k} \alpha^k - \alpha \\ &\geq \sum_{k=0}^{1} \binom{n}{k} \alpha^k - \alpha \\ &= 1 + n\alpha - \alpha \\ &\geq 1 + \alpha - \alpha \\ &= 1\,. \end{aligned}$$

(By the way, it's worth looking carefully at this derivation to see how it relies on the facts that $n \geq 1$ and $\alpha > 0$.) Finally, since f is a polynomial function, it is continuous everywhere; in particular, it is continuous on $[0, \alpha + 1]$.

So f is continuous on $[0, \alpha + 1]$, and $f(0) < 0 < f(\alpha + 1)$. Therefore, by IVT, there exists a number $c \in (0, \alpha + 1)$ (whence c is positive) such that $f(c) = 0$, that is, $c^n = \alpha$, QED.

This proof is adapted from Bowman's lecture notes for Math 117 (p. 69); see http://www.math.ualberta.ca/~bowman/m117/m117.pdf. It has a few nice features.

First, it illustrates the idea of turning a problem of the type "there exists a number with such-and-such a property" into a problem of the type "such-and-such a function has a zero". This is a very common maneuver. (That's why high school math includes learning ways to find zeroes.)

Second, it illustrates how to use IVT to show the existence of a root even if you can't (or don't want to) compute it.

Third, where Bowman simply says that "$f(x) > 0$ for sufficiently large x" (which could be deduced immediately from the fact that $\lim_{x\to\infty} f(x) = \infty$), the proof above constructs an explicit x for which this is true, namely $\alpha + 1$. And the construction is, in my not-so-humble opinion, pretty slick.

(It's not clear that slickness is a desirable property in proofs, but I take a certain half-guilty pleasure in it. A more natural approach to that construction is to consider that the only number we know about so far is α, so we might well wonder whether $0 < f(\alpha)$, that is, whether $\alpha < \alpha^n$. If $n \geq 2$ and $\alpha > 1$, then this is true. Then we have to deal with the remaining cases: if $n = 1$, then just short-circuit the whole proof by taking $c = \alpha$; if $n \geq 2$ and $\alpha < 1$, then $0 < f(1)$; if $n \geq 2$ and $\alpha = 1$, then $0 < f(2)$. Taking $\alpha + 1$ avoids this whole disgusting mess of cases.)

The disadvantage of this proof is that you have to develop a fair bit of theory before you can give it. First you need the theory of limits of sequences; then the theory of limits of functions; then the theory of continuity, which includes IVT. It's a long way from the axioms for $\mathbb{R}$ to IVT, or even from regular algebra to IVT. The existence of nth roots feels like a fundamental fact about the real numbers; it should be possible to prove it in a more elementary way.

That brings us to…

2.2 Proof by Rudin

(Rudin wrote *Principles of Mathematical Analysis* back in the 1970s, a famous textbook for undergrad analysis courses.)

For concreteness, I'll give (most of) the proof for the special case $n = 3$. Everything done can be generalized.

Let α be a positive real number, and let n be a positive integer. We wish to show that there exists a positive number c such that $c^3 = \alpha$.

Let $S = \{x \colon x^3 < \alpha\}$. This set is bounded above (by $\alpha + 1$, for example), so by the completeness axiom it has a least upper bound. Let $c = \sup S$. (Note that, since $0^3 = 0 < \alpha$, we know that $0 \in S$, and so $0 \leq c$.)

Now, suppose $c^3 < \alpha$. Choose a positive number h such that

$$h < 1 \qquad \text{and} \tag{1}$$

$$h < \frac{\alpha - c^3}{3c^2 + 3c + 1} . \tag{2}$$

(It is possible to choose a positive h satisfying the second inequality because the fraction is positive: $0 < \alpha - c^3$ by hypothesis, and $0 < 3c^2 + 3c + 1$ because $0 \le c$.) Then

$$\begin{aligned}(c+h)^3 &= c^3 + 3c^2h + 3ch^2 + h^3 \\ &= c^3 + h(3c^2 + 3ch + h^2) \\ &< c^3 + h(3c^2 + 3c + 1) && \text{by (1)} \\ &< c^3 + \alpha - c^3 && \text{by (2)} \\ &= \alpha .\end{aligned}$$

Therefore $c + h \in S$. But since h is positive, $c < c + h$, so this contradicts the fact that c is the supremum of S. Thus our supposition that $c^3 < \alpha$ is false.

A similar proof shows that $c^3 > \alpha$ is also impossible. (Choose the same h, except with the numerator reversed, and show that $(c - h)^3 > \alpha$. But then $c - h$ is an upper bound for S, lesser than the least upper bound c.)

Therefore $c^3 = \alpha$, as desired.

This proof is quite ingenious. As promised, it's close to the axioms; there is no appeal to a full-blown theory of limits and continuity and whatnot.

3 Inequality

We looked again at the arithmetic-geometric mean inequality and its application to maximizing the area of a rectangle. (See the notes for May 9.)

In Math 225 we looked (very briefly) at the Cauchy-Schwarz inequality, which states that in any inner product space we have

$$|\vec{u} \cdot \vec{v}| \le \|\vec{u}\| \|\vec{v}\| ,$$

and equality is attained exactly when $\vec{u}$ and $\vec{v}$ are proportional.

An application: What is the distance from the origin to the plane with equation $2x + 3y + 6z = 7$? By distance, of course, we mean the least distance; that is, we wish to find the minimum possible value of $\sqrt{x^2 + y^2 + z^2}$, subject to the constraint that $2x + 3y + 6z = 7$. Applying Cauchy-Schwarz to $\mathbb{R}^3$ with the usual dot product and norm, we see that

$$2x + 3y + 6z \le \sqrt{x^2 + y^2 + z^2}\sqrt{2^2 + 3^2 + 6^2} ,$$

and equality is attained when (x, y, z) and $(2, 3, 6)$ are proportional. Thus

$$\sqrt{x^2 + y^2 + z^2} \geq \frac{2x + 3y + 6z}{\sqrt{2^2 + 3^2 + 6^2}} = \frac{7}{\sqrt{2^2 + 3^2 + 6^2}} = 1 ,$$

and equality is attained for suitable (x, y, z). Therefore the distance from the origin to this plane is 1.

Once we've stated the problem as one of minimizing a norm, keeping a dot product constant (or maximizing a dot product, keeping a norm constant), the Cauchy-Schwarz inequality should leap to mind — it relates norms to dot products. Similarly, the arithmetic-geometric mean inequality should leap to mind when faced with a problem of minimizing a sum, keeping a product constant (or maximizing a product, keeping a sum constant).